Cruz *et al*. 10.1073/pnas.0509276102.

Supporting Information

Files in this Data Supplement:

Supporting Text
Supporting Table 3
Supporting Figure 4
Supporting Figure 5
Supporting Figure 6
Supporting Table 4

Supporting Figure 4

Fig. 4. Secondary structure propensity per amino acid for each of the trajectories. A turn in the sequence is denoted by a solid line, an -helix by a dotted line, a -helix by a slash-dotted line, and a coil (none of the above) by a dashed line.

Supporting Figure 5

Fig. 5. Scalar product *D* between the C -C vector of Val-24 and the C -C vector of Lys-28 for each trajectory. The plots at the right of each graph are their normalized histograms (*x* axis is vertical). The scales of the axes are the same for all of the histograms.

Supporting Figure 6

Fig. 6. Index (*Upper*) and instantaneous number (*Lower*) of intrapeptide HBs for each trajectory as a function of simulation time. Different HB indices in the *y* axis in the upper plots correspond to different HBs (small solid square).

Supporting Text

Thermodynamic Quantities. In Table 3 we list the thermodynamic quantities for all of the trajectories. These data were collected during the production runs by using the NVT ensemble rescaling velocities every 2 ps to maintain the temperature at $T = 283$ K. The thermodynamic quantities in Table 3 are very similar for trajectories [RC], [DU], and [RCS] where the sizes of the boxes and mass densities are very close to each other. For trajectories [P1] and [P2] the increased size of the box results in a reduced mass density of ~7.5% when compared with the other trajectories. Because the trajectories use the NVT ensemble, a bigger box size results in a negative pressure for trajectories [P1] and [P2] of about -142 MPa each.

Secondary Structure. In Fig. 4 we show the (total) secondary structure propensity per amino acid for each trajectory, calculated by using the data from Fig. 2, where each data point is the percentage of time that each amino acid spends in a given secondary structure. In all trajectories, except trajectory [DU], the turn propensity (Fig. 4, solid line) shows a cusp-like curve (usually around Gly-25–Lys-28) that increases for the central amino acids and decreases monotonically approaching either terminus. Specifically, trajectory [RC] shows a Gly-25–Lys-28 turn caused by the low $R(4,8)$ values placing Val-24 and Lys-28 in close proximity {Fig. 1 [RC] and high event time of $R^*(4,8)$ in Table 1}. Trajectory [P1] exhibits an additional -helix propensity involving the Glu-22–Ser-26

amino acids. In trajectory [P2] the center of the turn is primarily situated on the other side of the chain, Glu-22–Ser-26. In trajectory [RCS], the turn propensity within the sequence Val-24–Lys-28 is high because of a high number of $R^*(4,8)$ events in consort with a high increase of $R^*(2,8)$ and $R^*(3,8)$ (see Table 1).

Similarly, in trajectory [DU] the amino acids have a high propensity for a turn. However, the turn secondary structure seems to be split into two parts; the Ala-21–Val-24 part and the Ser-26–Gly-29 part. The key difference responsible for the two-turn layout seems to be the lack of charge in the Gln-22 [resulting in fewer $R^*(2,8)$ events] that destroys coherence in the chain, thus splitting it into two regions at the Gly-25 amino acid.

Relative Orientation Between Val-24 and Lys-28. When the C atoms of Val-24 and Lys-28 are in

close proximity during hydrophobic events, a natural question is whether this close proximity has any effect on the dynamics of the side chains (SD) of Val-24 and Lys-28. For this, we assess the relative orientations of these two side chains by calculating the normalized scalar product D between the vectors formed by the side chains of Val-24 and Lys-28. Values for D will range from -1, if the two side chains point away from each other or toward each other, to +1, if the two side chains are parallel to each other. In Fig. 5 we show the time evolution of D for each trajectory with accompanying histograms to the right. By comparing Figs. 5 and 1 we observe a persistence of the relative orientation of the side chains of Val-24 and Lys-28 during extended events. For instance, in trajectories [RC], [P1], [P2], and [RCS] (Fig. 5) the $R^*(4,8)$ events are accompanied by a rather stable (in time) D event between simulation times of 60 and 90 ns in [RC], 20 and 50 ns in [P1], 75 and 80 ns in [P2], and 100 and 120 ns in [RCS].

In the normalized histograms of D (right of Fig. 5), we observe that for all trajectories except [DU] there are bumps or skewness in these histograms, a result of the persistent values of D during events. Trajectory [DU], however, does not show appreciable skewness or bumps having a rather flat

histogram as compared with the others. This result emphasizes the relative motility of the Val-24 and Lys-28 side chains as opposed to other trajectories in which a formation of the loop restricts movement of these side chains.

HBs. In Fig. 6 we show plots of the occurrence of different HB (*Upper*) and instantaneous total number of HBs (*Lower*) as a function of time for HBs that form between atoms of the decapeptide. All of the SBs [*R**(2,8) and *R**(3,8)] are included and counted as regular HBs. In the upper plots, different HBs are assigned a unique arbitrary index used in the *y* axis to differentiate between them. Horizontal lines in the upper plot illustrate how particular HBs are active during events. By comparing Figs. 6 and 1 we observe that the total number of HBs (Fig. 6, lower plots) typically increases in the range from one to four HBs when particular folding events occur. For instance, the longest *R**(4,8) event in Fig. 1 [RC] is accompanied by an increase of about one HB in Fig. 6 [RC]. Similarly, the helix event in [P1] involves formation of at least four HBs. In the same manner R_g* events are correlated with an increase in the number of HBs.

To determine the identity of the HBs occuring during hydrophobic events, we list in Table 4 the HBs along with the HB count per trajectory. In this table, HB counts amounting to 1% of their trajectory totals are not listed. HBs are grouped according to those: (*i*) forming SBs, (*ii*) between a charged side-chain and a noncharged atom (Charged-SD), (*iii*) between backbone atoms (Backbone), (*iv*) between backbone and a side-chain atom (SD-Backbone), and (*v*) between one of the termini and another atom (N-C-Terminal HBs). The line at the bottom of each classification indicates its percentage relative to the total number of HBs in that trajectory. These percentage results are meant to be compared across trajectories while the absolute numbers for the HBs counts can only be compared within a single trajectory. The three most important HBs are marked in bold.

In Table 4, by grouping the HBs by the location of the atoms in these bonds, we can establish the correspondence between monomer structure and HBs. In most trajectories, the biggest percentage of HBs is contained within the Backbone HBs classification (except [RCS]). However, because the resulting secondary structure is different for each trajectory (Fig. 2) the role of HBs is also different in each trajectory. In trajectory [RC] the backbone HBs plus a high count of SB HBs give this monomer a tight stable loop. In trajectory [P1], the backbone HBs between Ala-21(O)-Ser-26(H), Glu-22(O)-Asn-27(H), and Asp-23(O)-Lys-28(H) contribute to the formation of the -helix. In trajectory [P2], its backbone HBs also contributes to a -helix with only one dominant HB between Val-24(O)-Gly-29(H). In trajectory [DU], although most of its HBs form within backbone HBs, a single persistent loop is not formed. The origin of the split turn in [DU] can be traced to the Lys-28(O)-Gly-25(H) and the Val-24(O)-Ala-21(H) HBs. Finally, [RCS] has the most HBs forming SBs with a minority of backbone HBs, making this trajectory the one with the most stable loop of all.